

AMENDMENT TO THE CLAIMS

1. (Cancelled)
2. (Cancelled)
3. (Cancelled)
4. (Cancelled)
5. (Cancelled)
6. (Cancelled)
7. (Cancelled)
8. (Cancelled)
9. (Cancelled)
10. (Cancelled)
11. (Cancelled)
12. (Cancelled)
13. (Cancelled)
14. (Cancelled)
15. (Cancelled)
16. (Cancelled)
17. (Cancelled)

18. (Cancelled)

19. (Cancelled)

20. (Cancelled)

21. (Cancelled)

22. (Cancelled)

23. (Currently Amended) A method of selecting speech segments for concatenative speech synthesis, the method comprising:

parsing an input text into speech units;

identifying context information for each speech unit based on its location in the input text and at least one neighboring speech unit;

identifying a set of candidate speech segments for each speech unit based on the context information, wherein identifying a set of candidate speech segments for a speech unit comprises applying the context information for a speech unit to a decision tree to identify a leaf node containing candidate speech segments for the speech unit, wherein identifying the sequence of speech segments comprises using an objective measure comprising one or more first order components from a set of factors comprising:

an indication of a position of a speech unit in a phrase;

an indication of a position of a speech unit in a word;

an indication of a category for a phoneme preceding a speech unit;

an indication of a category for a phoneme following a speech unit;

an indication of a category for tonal identity of

the current speech unit;
an indication of a category for tonal identity of
a preceding speech unit;
an indication of a category for tonal identity of
a following speech unit;
an indication of a level of stress of a speech
unit;
an indication of a coupling degree of pitch,
duration and/or energy with a neighboring unit;
and
an indication of a degree of spectral mismatch
with a neighboring speech unit, and;—and

identifying a sequence of speech segments from the
 candidate speech segments based in part on a
 smoothness cost between the speech segments; and
generating synthesized speech using the sequence of
speech segments without further prosody
modification.

24. (Cancelled)

25. (Currently Amended) The method of claim 24—23 wherein identifying a set of candidate speech segments further comprises pruning some speech segments from a leaf node based on differences between the context information of the speech unit from the input text and context information associated with the speech segments.

26. (Original) The method of claim 23 wherein identifying a sequence of speech segments comprises using a smoothness cost that is based on whether two neighboring candidate speech segments appeared next to each other in a training corpus.

27. (Cancelled)

28. (Currently Amended) A method of selecting speech segments

for concatenative speech synthesis, the method comprising:

parsing an input text into speech units;

identifying context information for each speech unit based on its location in the input text and at least one neighboring speech unit;

identifying a set of candidate speech segments for each speech unit based on the context information, wherein identifying a set of candidate speech segments for a speech unit comprises applying the context information for a speech unit to a decision tree to identify a leaf node containing candidate speech segments for the speech unit,

The method of claim 23 wherein identifying a the sequence of speech segments comprises using an objective measure comprising one or more higher order components being combinations of at least two factors from a set of factors including:

an indication of a position of a speech unit in a phrase;

an indication of a position of a speech unit in a word;

an indication of a category for a phoneme preceding a speech unit;

an indication of a category for a phoneme following a speech unit;

an indication of a category for tonal identity of the current speech unit;

an indication of a category for tonal identity of a preceding speech unit;

an indication of a category for tonal identity of a following speech unit;

an indication of a level of stress of a speech unit;

an indication of a coupling degree of pitch, duration and/or energy with a neighboring unit;
and

an indication of a degree of spectral mismatch
 with a neighboring speech unit;
identifying a sequence of speech segments from the
candidate speech segments based in part on a
smoothness cost between the speech segments; and
generating synthesized speech using the sequence of
speech segments without further prosody
modification.

29. (Currently Amended) The method of claim ~~24~~—28 wherein identifying a sequence of speech segments further comprises identifying the sequence based in part on differences between context information for the speech unit of the input text and context information associated with a candidate speech segment.

30. (Cancelled)

31. (New) The method of claim 28 wherein identifying a set of candidate speech segments further comprises pruning some speech segments from a leaf node based on differences between the context information of the speech unit from the input text and context information associated with the speech segments.

32. (New) The method of claim 28 wherein identifying a sequence of speech segments comprises using a smoothness cost that is based on whether two neighboring candidate speech segments appeared next to each other in a training corpus.